Explorations of Display Interfaces for Virtual Reality

Michael F. Deering
Sun Microsystems Computer Corporation
2550 Garcia Avenue, Mountain View, CA 94043
e-mail address: michael.deering@Eng.Sun.COM

ABSTRACT

Three alternatives to the traditional head mounted Virtual Reality display are described. One configuration is the Virtual Holographic Workstation; an external stereo CRT viewed by a user wearing head tracking stereo shutter glasses. Another is a Computer Augmented Reality camcorder, where virtual objects are composited onto live video using six axis tracking information about the location of the video camera. The third system is the Virtual Portal; an entire room is turned into a high resolution inclusive display by replacing three walls with floor to ceiling rear projection stereo displays. Details of the systems and experiences and limitations in their use are discussed.

1.0 INTRODUCTION

Virtual Reality has become almost synonymous with bulky head mounted displays and gloves. Certainly such gear makes the point that this is a different human-computer interface technology than most people have experienced before. But what is really important about the visual interface is not the bulky paraphernalia, but the more precise generation of a number of visual cues that previous technology has lacked.

There are other methods of generating these visual cues, and interesting sub-combinations of such cues, that also merit attention. These alternatives have practical advantages both for commercial applications and VR research. This paper will describe several such alternative systems built at Sun Microsystems, and some of the experience gained in using them.

2.0 VISUAL CUES

All synthetic visual display systems work by providing a number of natural visual cues in their imagery. Most of these cues have also been produced by painters for hundreds of years: shape, texture, depth occlusion, haze, etc. Head mounted displays systems combine three additional powerful visual cues that have been lacking in most widely used previous technologies:

1. They are head coupled. The view is directly controlled by six axis motion of the users head, the same as the user's vision works in the physical world (and different than most all prior computer displays have worked).
2. They generate stereo imagery with correct perspective. Each eye gets a different view, as determined by the individual user's eye separation (intraocular distance) in space.
3. The imagery is inclusive. The computer generated images subtends at least 65° of the viewers field of view, enough to give the view a sense of "presence".

Because of technical limitations, many past VR systems have *not* supported a fourth important cue that has been a recognized necessary basis for synthetic display since the 1820's:

4. The view changes in real time, at or beyond the threshold of the human visual system's motion fusion rate.

0-7803-1363-1/93 $3.00 © 1993 IEEE

Various degrees of inclusion have been attempted by historical cinema systems; as has stereo display, though always with some amount of distortion in the perspective. The really new ingredient is dynamic head tracking. Even most commercial flight simulators historically have not directly used head position information. The only other display technology where changes in user head position produce different stereo imagery gave been holography and various forms of "solid" display, none of which have been capable of dynamic update of scene contents. Something different happens when all four of these effects are combined, and this combined effect in VR systems is what people are excited about.

Recent perceptual work is showing that the combination of head coupled display and stereo is very much more powerful than just stereo by itself [Arthur93]. This is because the human visual systems creation of 3D perception does not rely just on static stereo cues. Rather, the dynamic motion of the viewer's head provides a rich continuous stream of new cues about shape and distance in the real world. Indeed, one-eyed virtual displays employing headtracking alone *without* stereo, when all physical viewing parameters are adjusted properly, can provide a stronger sense of 3D than static 3D displays.

> *Experiment:* Assuming that you have normal binocular vision, close one eye and walk around a complex 3D object. Does it feel more real than viewing a static stereo photograph of the same object would?

The correct visual stimuli is almost never generated by stereo image pair displays if headtracking is not used. This is because even a stereo image pair *assumes* a single, precise, fixed head position of the viewer in space (not to mention an assumed interocular distance). Only if the viewer's head is at this one single position and orientation will the perspective and stereo information be correct. Without headtracking the best one can do is to use some fixed viewing apparatus.

3.0 ALTERNATIVES TO HEAD MOUNTED DISPLAYS

We have built three VR display alternatives to the head-mount at Sun Microsystems.

Virtual Holographic Workstation. Using a conventional desktop CRT this combines cues 1, 2, and 4 to produce a head tracked stereo display, but without inclusion.

Cameye. This is a Computer Augmented Reality video camera. It combines 1, 4, and part of 2 as an overlay on the physical reality to create stabilized virtual objects mixed in with the real world.

The Virtual Portal. Three rear screen stereo displays provide all the cues, but without the weight and low resolution of existing headmounted displays.

Each of these systems offers some advantages over a traditional head mount display system, traded off against some limitation.

4.0 VIRTUAL HOLOGRAPHIC WORKSTATION

Combining a 3D workstation, head-tracked stereo shutter glasses, a stereo CRT, and a 3D mouse results in the Virtual Holographic Workstation. Mathematical details are described in [Deering92]. [Paley92] is a similar system. By updating the screen with stereo pairs of 3D shaded images in sync with the user's head motion, the results are stabilized solid appearing 3D objects in space. With the 3D mouse, an application can process precise 3D input events from the user, knowing that the user's visual perception of the location of the virtual objects is precisely regis-

tered with the user's perception of near-by physical objects, like their hands and a work tip attached to the 3D mouse.

The accuracy is so high that one can hold up a physical ruler against the virtual objects to measure them. In one experiment, the user can check the fit of two mechanical parts together — one virtual, the other physical —and visually detect tolerance errors as small as a few mils.

The Virtual Holographic Workstation trades off inclusion for resolution and practicality. The user need only wear a lightweight pair of headtracking stereo shutter glasses. Because it uses a conventional 21" CRT, the virtual display only subtends on the order of 45° of the users field of view, which is substantially less than the threshold for inclusion. Therefore the effect produced is not of the viewer being immersed in a totally virtual environment, but rather of virtual objects inhabiting a limited portion of our physical world. The resolution is considerably higher than that of most all head mounted displays, as with 960 pixels spread across a 45° field of view, the resolution is on the order of 2.8 minutes of arc per pixel. When antialiased lines are being used, the effective resolution is enhanced by nearly a factor of three, to 56 seconds of arc per pixel, which is getting close to the maximum human foveal resolution of approximately 30 seconds of arc.

A key question for human perceptual studies of Virtual Reality to investigate is "how real is virtual reality?" By this we mean how well can visual tasks previously performed in the physical world be performed in the virtual, or mixed physical and virtual world? How fast? How accurate? But it is hard to address these questions with the extremely low resolutions and optical distortions of most present head mounted displays. The Virtual Holographic Workstation lets us side-step such limits.

Because the system also functions as a normal workstation, running a normal window system, and physically looks like a conventional desktop workstation (except for the triangular ultrasonic transmitter perched on top of the CRT), for many environments that configuration is more practical than an inclusive headmounted display.

5.0 COMPUTER AUGMENTED REALITY CAMERA

Rather than locking the user up in a totally virtual world, computer augmented reality allows virtual objects to be mixed in with the real world. Typically in the past this has been done by half silvered mirrors combining computer imagery (typically wire frame) with a direct physical view. The problem with this approach is that the virtual objects always appear partially transparent, even when drawn as solids. We experimented with the alternative of using video keying to insert real time 3D shaded images on top of live video. A 6D tracking device was attached to the video camera taking the images. In this manner the user holding the camera can pan it about the physical environment, observing virtual objects just as if they were any other object within the camera's view (figure 1). Our camera is actually a modified camcorder, with an integral color LCD view-finder. When the viewer is holding the camera to their eye, the somatic cues about the viewer's own position and orientation in the environment add to the sense of reality of the image. In this sense, the system is effectively a one-eyed see-through head mounted display.

Computer Augmented Reality has many potential industrial and medical applications. Some applications operate by superimposing known or sensed data about the internal details of a physical object directly on that object. Examples include superposition of medical data on patients [Bajura92], superposition of technical data on machinery [Feiner92], superposition of plumbing, electrical, and structural data on buildings, and superposition of stress and strain data on mechan-

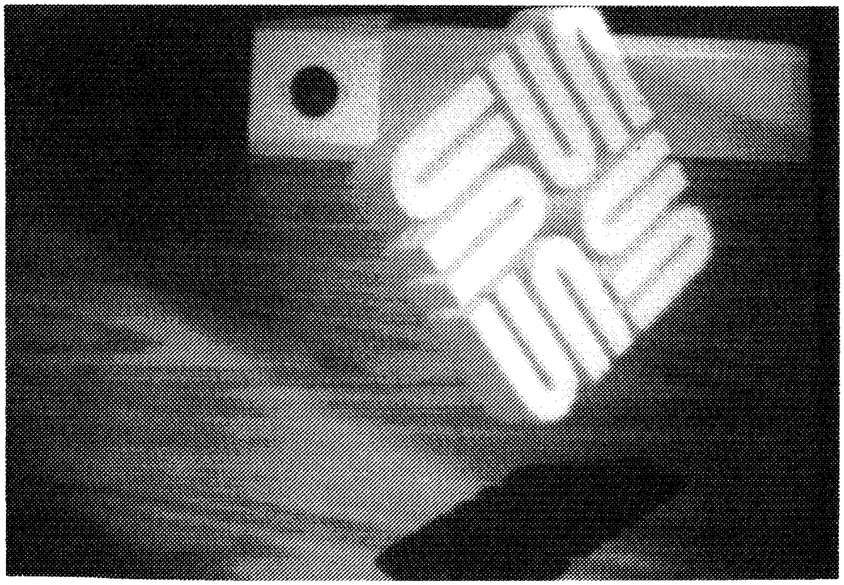

Figure 1. Computer Augmented Reality. Virtual logo and shadow on real desktop, shot live.

ical objects. Other Computer Augmented Reality applications involve superimposing not yet existent objects onto the real world. Examples include showing possible office partitions in a bare office, showing the next days/weeks planned building construction superimposed on top of construction in progress, and superimposing new clothing on a customer.

We have done some experimenting with this last example, of showing a person wearing something virtual. To do this, six axis information about the location and orientation of the users body parts is needed. To start simply, we prototyped a system that just tracked the users head, since we already had headtracking experience and technology. Several virtual objects were constructed for the user to wear: a Ronald Reagan mask (with eyeballs), a spinning propeller beanie, a halo, a cartoon thought bubble (complete with witty sayings), and a fire-breathing mode. The resulting live images were very natural, the high speed head tracking and accurate perspective caused the virtual object's position and motion to stay in nearly exact sync with the user's head, to which they appeared convincingly attached.

As our video input is based on NTSC, the entire system runs at NTSC resolution. In this mode its pixel resolution is limited to 640×480, with a camera imposed narrow field of view of 40°. On the built-in view finder, the resolution is even more severally constrained, as its color LCD display is the same low resolution device as is used in many head mounted displays.

But the feeling of presence is not as severely diminished as a total virtual world would be at this resolution. This is because most of the world is the *real* world around the user, virtual objects exist only at limited locations embedded in the physical world. Effectively the user can "foveate" a particular virtual object by pointing the camera at it, without losing the frame of reference of the rest of the physical room or space about him.

As others have noticed with video super-positioning systems, occlusion is a very strong depth cue, and will override most all others. This means that no matter how correct a virtual object looks to be

positioned within the physical room, if something clearly closer than it pass "in front" of the virtual object, but the virtual object is *not* occluded, the virtual object perceptually appears to shrink in size and distance, so as to appear to still be in front of the physical object. (A common example is someone placing their hand very near the camera.) This problem can be overcome if accurate tracking information is known about static or mobile physical objects that might occlude virtual objects. Thus it is possible to construct in virtual space an invisible analog of the physical object, but leaving proper distances in the Z-buffer. In this way virtual objects that need to be behind physical objects will be properly occluded by the Z-buffer rendering process. More complex alpha channel techniques are required if one wishes to mix partially transparent physical and virtual objects.

Angular accuracy is much more important in this system than the others we describe in this paper. Small changes in the angle of the camcorder need to cause large shifts in the screen space position of virtual objects supposedly located several feet in front of the camera. With our $40°$ horizontal field of view, and 640 pixels across, a shift of only 4 minutes of arc should cause a virtual object to shift by a pixel. This is below the rated angular accuracy of our position and orientation tracker. Computer Augmented Reality applications will need to employ higher accuracy trackers to achieve reasonable stabilized virtual objects.

6.0 THE VIRTUAL PORTAL

The Virtual Portal is a full inclusive VR display system without the weight, distortion, and resolution limitations of traditional head mounted displays [Deering93]. The approach is to turn all visible surfaces into display surfaces. The Virtual Portal is a small room, where three of the walls are actually floor to ceiling rear projection screens. Three graphics workstations drive three projection CRT's, using genlocked field sequential stereo (figure 2). The single viewer walks about in the room wearing stereo head tracked glasses, with a six axis headtracker mounted in the ceiling. (The Cave is a similar system [Cruz-Neira92]).

To the user, the three walls (and their seam) melt away, producing a fully inclusive 3D display. Virtual objects can exist at distance ranges from kilometers away to right inside the room, just tens of centimeters away from the viewer. Because of the very wide field of view, the viewer perceives herself as fully immersed in the virtual environment. Each screen has a resolution of 680×960 pixels (for each eye), a total of 2040×960 stereo color pixels. The lightweight stereo shutter glasses support a very wide field of view: $95°$ in the horizontal direction with $75°$ stereo overlap. The pixel solid angle varies with user head position, assuming head placement at the center of the room, the resolution is approximately 8 minutes of arc per pixel. Antialiasing can increase the effective resolution to better than 3 minutes of arc for feature placement.

In contrast, many traditional LCD based head mounted displays have only 208×138 pixels per eye, across a visual field of $65°$, resulting in a resolution of 20 minutes of arc per pixel. Also, current LCD displays have limited greyscale range, reducing the ability to simulate higher resolutions through antialiasing.

Even more importantly, the Virtual Portal's nearly perfectly flat screens have none of the tremendous optical distortions present in most current head mounted displays, regardless of resolution. Human stereo perception is quite sensitive, the elimination of such distortions results in a much greater and more natural feeling of presence. Head mounted display systems are also subject to other optical effects, such as transverse chromatic distortion, that are bypassed by the rear-screen projection approach.

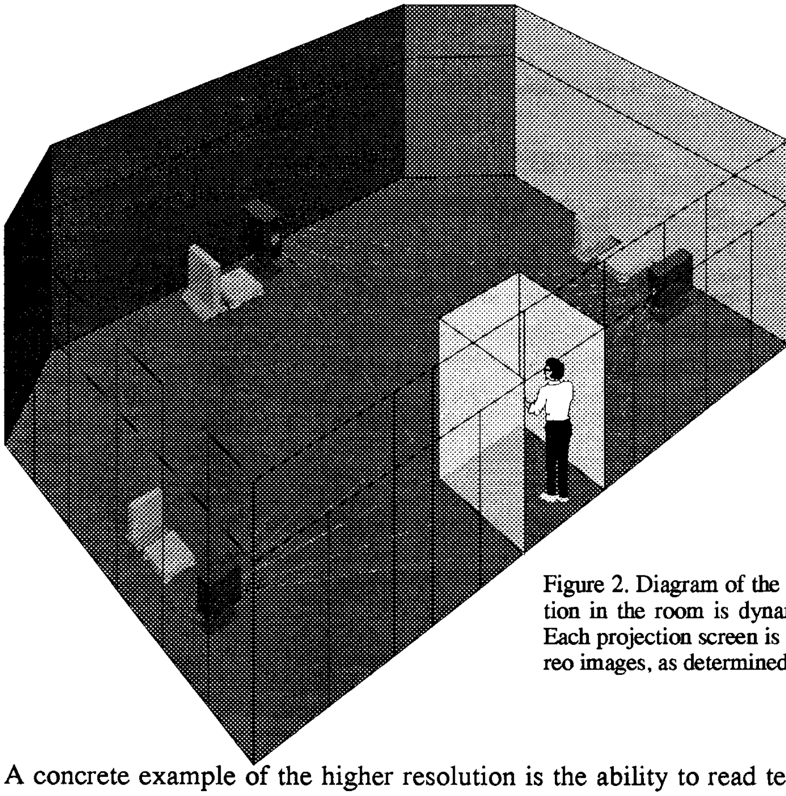

Figure 2. Diagram of the Virtual Portal. The viewer's position in the room is dynamically tracked by the computer. Each projection screen is displaying real-time animated stereo images, as determined by the viewers current position.

A concrete example of the higher resolution is the ability to read text in virtual space. Both the Virtual Portal and the Virtual Holographic Workstation support a terminal emulator that appears as a rumpled sheet of paper in virtual space. The text is generated by high quality antialiased vector fonts, rendered just in front of the paper background. The text is highly readable, at a full 80 column width. This should be contrasted with the lack of readable text in most head mounted display systems.

Our experience so far is that, like head mounted displays, the Virtual Portal also generates a strong sense of "presence". Because of the higher effective resolution and lower distortion, the stereo imagery is more stabilized. Subjectively many viewers felt that the over all experience was stronger than with existing head mounted displays, but there are not yet any clear objective ways to demonstrate this, other than the ability to read text, and for viewers to repeatedly, accurately visually estimate distances and sizes of virtual objects.

As a practical technology, the applications for the Virtual Portal are constrained to only a few specialized ones by the large amounts of floor space required. The motivation for building the Virtual Portal was more for experimentation purposes. Someday display technology will advance enough to provide adequate head mounted displays: resolutions of 2K×2K or greater, full color (with a reasonable number of grey scale levels and contrast), true non-distorted (or electronically corrected) optics (including correction of transverse chromatic distortion), light weight (for long wearing), and at reasonable costs. Until then it is very hard to even experiment with Virtual Reality applications, when you can't even read text. The Virtual portal thus functions as a very large stand-in prototype of these future mass market head mounted displays. It allows us to experiment with what sort of software support and applications will be viable with such future displays ahead

of time. To this end, we are conducting a number of visual experiments with the Virtual Portal display, as well as developing our VR software support environment.

7.0 PHYSCIAL CONSTRUCTION

All three systems use Sun SPARCstation 2GT's for image generation and computing. All 3D six axis tracking is performed by Logitech ultrasonic trackers.

The Virtual Portal's three rear projection screens were formed by wrapping a single flexible screen around a large frame. Each screen is driven by an Electrohome ECP4100 video projector. Special high speed green phosphor tubes from StereoGraphics eliminate ghosting of one eye's image into the other. The projectors are mounted sideways to shorten the projection distance required to achieve a floor to ceiling display. Video is genlocked to the stereo left frame.

8.0 CONCLUSIONS

Three alternative approaches to Virtual Reality display have been presented. Each overcomes a particular limitation of traditional approaches, extending our ability to experiment with Virtual Reality applications.

9.0 ACKNOWLEDGMENTS

The author would like to thank Keith Hargrove, Michael Neilly, Bob Cannataro, and Scott Nelson for their technical work in building the systems described here.

10.0 REFERENCES

[Arthur93] Arthur, Kevin, Kelly Booth, and Collin Ware. 3D Task Performance in Fish Tank Virtual Worlds, to appear in *ACM Transactions on Information Systems*, Special Issue on Virtual Worlds, July 1993.

[Bajura92] Bajura, Michael, H. Fuchs, and R. Ohbuchi. Merging Virtual Objects with the Real World: Seeing Ultrasound Imagery within the Patient. Proceedings of SIGGRAPH '92 (Chicago, Ill, July 26-31, 1992). In *Computer Graphics* 26, 2 (July 1992), 203-210.

[Cruz-Neira92] Cruz-Neira, Carolina, D, Sandin, T. DeFanti, and J. Hart. The Cave: Audio Visual Experience Automatic Virtual Environment. In *Communications of the AC*, 35, 2 (June 1992), 64-72.

[Deering92] Deering, Michael. High Resolution Virtual Reality. Proceedings of SIGGRAPH '92 (Chicago, Ill, July 26-31, 1992). In *Computer Graphics* 26, 2 (July 1992), 195-202.

[Deering93] Deering, Michael. Making Virtual Reality More Real: Experience with the Virtual Portal. In *Proceedings of Graphics Interface'93* (Toronto, Ontario, May 19-21, 1993), 219-226.

[Feiner92] Feiner, Steven, Blair MacIntyre, and Dorée Seligmann. Annotating the Real World with Knowledge-Based Graphics on a See-Through Head-Mounted Display. In *Proceedings Graphics Interface '92* (Vancouver, British Columbia, May 11-15, 1992), 78-85.

[Paley92] Paley, W. Bradford. Head-Tracking Stereo Display: Experiments and Applications. Stereoscopic Displays and Applications III (San Jose, California, February 12-13, 1992.). In *Proceedings of the SPIE* 1669, 1992.